Faculty Working Papers

MARKET STRUCTURE, INNOVATION, AND OPTIMAL PATENT LIFE

Lawrence M. DeBrock, Lecturer, Department of Economics

#659

College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

FACULTY WORKING PAPERS

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

March 25, 1980

MARKET STRUCTURE, INNOVATION, AND OPTIMAL PATENT LIFE

Lawrence M. DeBrock, Lecturer, Department of Economics

#659

Summary

This paper examines the question of optimal patent duration under different conditions of rivalry in the market for invention. As in previous studies on this subject, a static technological environment is assumed. Unlike earlier studies, however, the present paper permits the patent winner to further develop its invention if such additional expenditures are privately desirable. The results show that invention markets characterized by rivalry have shorter optimal patent lives than markets with monopoly inventors, a result counter to the findings of previous studies where all project specific R&D expenditures stopped upon conferral of the patent.

It is well known that the growth of inputs cannot, alone, explain overall output growth. A significant factor in explaining output growth is technical advance. Production functions are constructs that rest on an assumption of given technology; technical advance results in attractive shifts in most production functions. For many years, economists relegated technology as a parameter and thus examined the performance of economic systems by resource allocation among the variables of land, labor and capital.

The lack of attention to the production of new technology was somewhat understandable; technical improvements are essentially new knowledge and the price system has well known problems in dealing efficiently with such a "commodity." Technical advances are, for the most part, costly to arrive at but inexpensive to duplicate.

Findings by such authors as Schmookler (1966) pointed out the lack of good explanatory power the theory of the firm had with respect to technical advance; profit making firms were undertaking sizable amounts of research and development. Such allocation of resources toward something previously modeled as a parameter pointed to the need for more developed theory.

Much work has been directed at determining the relative efficiency of various market structures in the allocation of resources toward invention. It is not the direct purpose of this paper to make such examinations. An extensive survey of these studies has been provided by Kamien and Schwartz (1975). While the conclusions are not entirely clear, it seems certain that no single market structure results in a Pareto efficient allocation of resources to research and development (R&D).

In light of this, certain "second best" procedures must be examined. Social controls may provide an attractive method toward lessening the inefficiencies described above. While government subsidies for R&D and award programs exist, the most prevalent form of social control is the patent system. By its very nature, a patent brings monopoly problems into the allocation of resources but the arguments for turning to a second best solution have already been given. Awarding the inventor a monopoly on the use of his discovery for a specified period of time eliminates the free rider problem, thus restoring the incentive to invent. Extension of the duration of protection will increase incentives for private resource allocation into technical advance. Unfortunately, extension of patent protection definitionally brings with it the social inefficiencies recognized in a monopolistic market. The interaction of these two opposite forces implies a trade-off and the existence of an optimal patent life.

Studies of optimal patent life have not been numerous. Nordhaus (1969) examined the question with a formulation similar that first advanced by Arrow (1962). Assuming no competition in the inventive process, Nordhaus posits a concave production function relating inventive benefits to R&D expenditures, which he then uses to maximize a social welfare function. Scherer (1972) developed a partial geometric interpretation of the Nordhaus model, depicting private benefits and costs while neglecting to present the alterations that consumer surplus has on the geometry. Kamien and Scwartz (1974) bring in the realistic assumption of rivalry in the inventive process, and find a longer optimal patent life than

Nordhaus. However, their results are sensitive to two contestable assumptions: innovation stops at the awarding of the patent and the degree of rivalry is exogenously determined. The latter appears to be a methodological error; in a study directed at finding the optimal patent life, it is important to recognize that altering the length of patent protection will result in larger expected valve to the patentee and thus more intensive rivalry. It can be argued that the former assumption, that the invention innovation project size is a once for all decision, is unrealistic in the general sense.

Once the winning firm has been awarded the patent, it is not allowed to go on to develop the patent to its full potential; the invention model has a putty-clay nature with respect to R&D. Since each firm prior to the conferral of the patent chooses its project size according to an expected private benefit necessarily smaller than the actual private benefit of a winner it will choose a project size smaller than a monopoly inventor. Once the "winner" has been determined, it should act as a monopolist in that it sees the total benefit and cost, not the discounted benefits. In fact, experience has shown that this is the general case in real world innovation examples.

A useful analogy presents itself. In the classic Cournot model, as the number of firms increases, the individual firms output, q_i, decreases while market output, Q, increases from the monopoly level. In the Kamien and Schwartz model, increases in rivalry bring about decreasing project size for the individual firm; however, since there is only one winner who cannot develop its patent further due to the putty-clay nature of the model, the market project size also decreases.

Unlike the Cournot model, society sees no benefit from increased competition, only a diminished level of technical change.

The effect of rivalry then is to diminish the size of the improvement. In reality, while there may be some examples of putty-clay processes, we would expect the winner of a patent to continue development to the private optimum just as a Nordhaus monopoly inventor. As a result of their specifications, Kamien and Schwartz find the optimal patent life is longer when rivalry is present. Intuitively, since the degree of rivalry is exogenous, longer patent protection inflates private benefits somewhat offsetting rivalry's effect on project size, without altering the degree of rivalry and duplicative costs of the losers, which they incorrectly assume is exogenous to the model.

This paper will examine the issue of patent life and market struction using a different assumption. Rivalry will be specified as an endogenous function of the level of profits going to the winner. The invention-innovation scenario is a two stage process. The first stage is the competition to win the initial patent; the second stage allows the patent winner to continue to develop its invention just as would a monopoly inventor. Three conditions will be satisfied in arriving at the optimal patent life under rivalry: private profit maximization, social welfare maximization, and a zero expected profit equilibrium condition. We arrive at an optimal patent life shorter than the Nordhaus monopoly inventor model whereas Kamien and Schwartz found the optimal patent life to increase as one moves out of the monopoly situation.

Finally, we should note two other assumptions at the outset. The potential benefit, both private and social, of the inventor is assumed to

be of the "manna" type; that is, it suddenly appears and/or becomes known to the potential inventor. This is in contrast to models where the benefits are growing over time, or costs shrinking due to increased knowledge, making profitability a function of time. Such models will be examined in the subseqeunt essay. Additionally, timing factors are taken to be exogenous parameters in the present study. Granted, the timing of introduction decision is a key element in Kamien and Schwartz' analysis. However, they find (Kamien and Schwartz (1972), p. 10) that the effect of rivalry on timing is ambiguous. Additionally, Ippolito has shown that the direction of bias introduced by specifying timing factors exogenously is, a priori, indeterminant (p. 2). Thus, it does not appear that making such timing considerations endogenous would greatly alter the results.

Section I will present a generalized model of the Nordhaus monopoly inventor scenario with a more useful geometric interpretation than presently put forth. The second section will examine the effects of rivalry on the situation.

I. MONOPOLY INVENTOR MODEL

In this section we consider a market for invention with no rivalry; the firm proceeds, with certainty, toward some inventive goal secure in the knowledge that it has no competition for this goal. The firm sells a final product in a competitive market.[1] The invention is of the process improvement variety. While product inventions are not explicitly treated, one can imagine such invention as finally achieving costs low enough to warrant introduction of a product that has always been "demanded."

A. A Simple Model of Invention

Consider a market with demand and costs as shown in Figure 1. The private firm can invest in R&D expenditures to introduce a process improvement that lowers unit production costs. We assume that the effective R&D expenditures, I, are scaled in dollars so that the marginal cost of an extra unit of inventive input is a constant number, which we set at 1. That is, the variable I represents an efficient use of resources in its "production." We assume that the improvement is not so major as to result in an increase in quantity by a profit maximizing patent monopolist: the so called "run-of-the-mill invention" (Nordhaus, 1969). Improvements lower MC_0 to MC_1 in Figure 1, earning the patent holder returns of α as he charges royalty rate $MC_0 - MC_1$.[2]

Assume that α is an increasing concave function of the level of I and the size of demand, which we assume to be constant and suppress for the time being. A larger R&D project, i.e. larger I, will result in a more dramatic improvement. That is to say, a larger I will cause an even greater reduction in marginal production costs, which increases the area α in Figure 1. The inventor realizes a present discounted value of

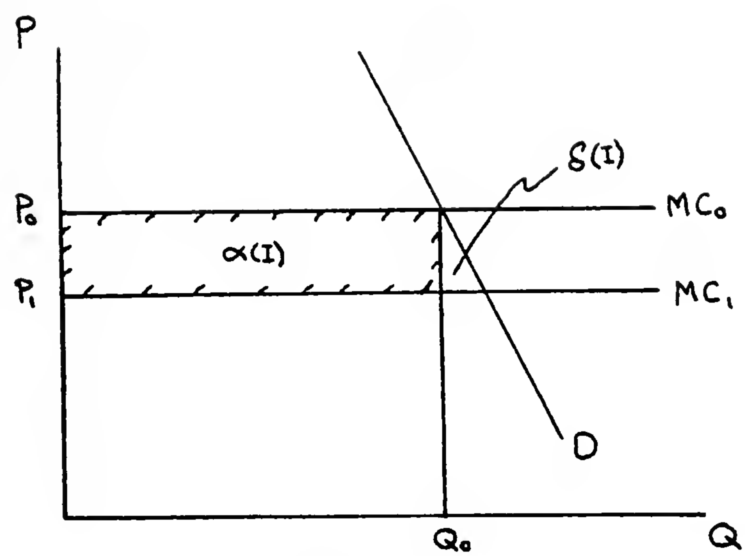


FIG 1.

$$V = \int_0^T \alpha(I)e^{-rt}dt - I \tag{1}$$

where r is the discount rate and T is the government administered patent protection period.[3] We assume that the patent holder licenses out the process at a negligible transaction cost, thus avoiding any adjustment costs and transition periods between the time he receives the patent and when he earns full monopoly rents.[4] Additionally, at time $T + \varepsilon$ the property rights to the invention become public and any firm can produce at MC_1. The users no longer pay royalties and competition drives the price down to MC_1.

Maximizing present discounted value by setting $dV/dI = 0$ gives

$$\alpha'(I) \int_0^T e^{-rt}dt - 1 = 0 \tag{2}$$

which can be solved for the profit maximizing level of R&D, I*. In Figure 2, this is represented by the intersection of the MC = 1 and the curve P at the level I*, here P is defined as the first term in (2).

Turning to the social benefits from invention, we need to examine the change in consumer surplus due to the introduction of the technical advance. By assumption, price does not change until time period T. Denoting the deadweight loss triangle as $\delta(I)$, where $\delta' > 0$ and $\delta'' < 0$ we have the change in consumer surplus

$$\Delta CS = \int_T^\infty (\alpha(I^*) + \delta(I^*))e^{-rt}dt$$

where the social discount rate is assumed to be equal to the private rate.

Finally, under the usual assumption, we will take social welfare as the efficiency measure of net surplus, W = PS + CS, recognizing the arguments sometimes voiced against such an assumption. Later it will be shown that changing the relative weights of CS and PS in W does not change the qualitative conclusions of the model. Recall that we are neglecting any possible external benefits of the patented invention and/or the patenting process itself.

Thus we have

$$W = \int_0^\infty \alpha(I^*)e^{-rt}dt + \int_T^\infty \delta(I^*)e^{-rt}dt - I^*. \tag{3}$$

This can be depicted in Figure 2 as the area under the curve S but above MC = 1, where s is the marginal social benefit curve:

$$S = \alpha'(I) \int_0^\infty e^{-rt}dt + \delta'(I) \int_T^\infty e^{-rt}dt \tag{4}$$

In Figure 2 we know that the change in consumer surplus due to invention is the area below S but above P while the additional present discounted value to the firm from inventive activity is the area below P above

marginal cost. Unconstrained, society would like to move out to the intersection of S with MC, but the fact that private firms do not recognize the additional consumer benefits from R&D leads to investment at a lower level. As mentioned in the introduction, we are forced into a second best situation.

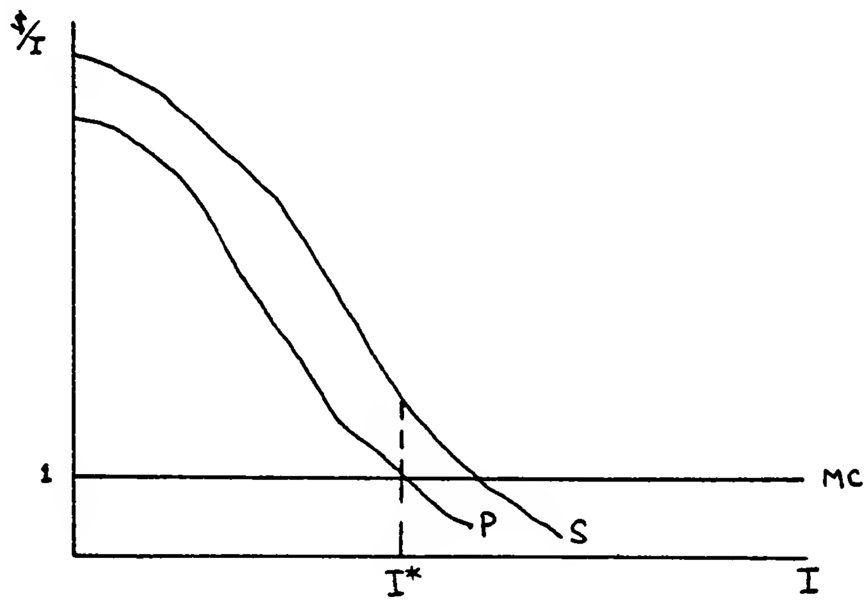


FIG. 2

Examination of the effects of a change in patent life is possible. If the patent life is shortened, T is decreased. From (2) and (4) it is obvious that such a decrease in T will lower P for any I while raising δ. A new, lower I* is arrived at, as shown in Figure 3. PS drops by the area A + B. However, CS increases by the area A + D - C. Net change

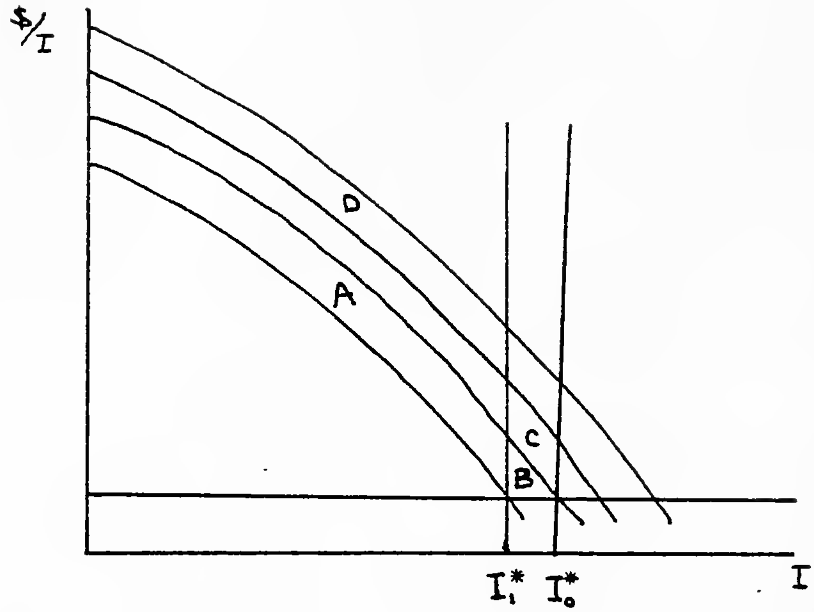


FIG. 3

in welfare is D - C - B; the area A is a loss to the private firm offset by a gain in consumer surplus. If this net change is positive, we should continue to shorten the duration of patent protection.

This simple model allows a more complete geometric interpretation of the Nordhaus model. Unlike Scherer's geometry, we can depict the social benefits and the second best nature of the I* arrived at even under optimal choice of patent life. However, the above model only shows the direction of beneficial changes in patent life. To better show the optimal T we now turn to a more general model.

B. <u>A Generalized Model of the Monopoly Inventory Case</u>

Consider an invention market with only one participant. The R&D investment will result in a process invention yielding present discounted value of extra profits in the production and sale of the industry's

output of R dollars when developed at time $t = 0$. The exact amount of this extra revenue depends upon the level of effective R&D expenditures, I, and on the duration of protection from some other firm(s) use of your idea/invention. In the present concern, this protection is in the form of a patent. However, we can imagine the given industry without the government protection device. Most likely, there will be a period of natural protection, $\hat{T}$, representing any imitation lag caused by such things as secrecy, etc. While the size of $\hat{T}$ will vary across industries, it is certainly bounded on the low side by zero. In any industry, if patent life is less than $\hat{T}$, then such a patent life is not relevant; no firm would take out a patent and reveal all the information as to its invention if the secrecy route led to a larger protection period. That is, optimal patent life is bounded by the value of $\hat{T}$ in the specific case, as will be shown later. For the present, assume that $\hat{T} = 0$. This will be relaxed shortly. Thus we have revenue from inventive activity:

$$R = R(I,T).$$

Specifically, we assume that R is an increasing concave function of the level of inputs

$$R_I > 0 \qquad R_{II} < 0 \qquad R_{II}R_{TT} - 2R_{IT} > 0$$

and at time $T + \varepsilon$ the extra revenues to the patent holder vanish as competitors drive the price of the industry product down to the new lower cost of production, as explained in part A.

The firm faces a cost schedule, C(I) which we assume to be of the general form

$$C' > 0, \quad C'' \geq 0.$$

Finally, assume the patent holder collects all returns to the patent over its lifetime; that is, imitation is not permitted. This is not unreasonable, given the monopoly position of the inventor in the first place.

The firm will maximize the present discounted value of the profit from inventive activity,

$$\pi = R(I,T) - C(I). \tag{5}$$

Using (5) it is possible to construct a reaction function for the firm, relating its profit maximizing level of I to any specific level of patent protection. Totally differentiating (4) and setting this equal to zero, we can solve for dI/dT which is the slope of an isoprofit contour in I, T space.

$$dI/dT = -\frac{R_T}{R_I - C_I} \tag{6}$$

If we call I* the privately determined profit maximizing level of I given some T we know the denominator satisfies

$$R_I - C_I \gtrless 0 \quad \text{as} \quad I \lessgtr I^*$$

Since the numerator of (6) is always positive, we know the slope of the isoprofit curves is

$$\frac{dI}{dT} \lessgtr 0 \quad \text{as} \quad I \lessgtr I^*$$

A representative isoprofit map is depicted in Figure 4. The reaction function is the locus of vertical slope points on the isoprofit map. From (5) we know that $R_I - C_I = 0$ if the firm is maximizing profits. As T is varied, this maximization condition determines how the firm must alter I. This gives a "demand" curve relationship of I to T. Differentiating this with respect to T we can find the slope of the reaction function is positive. Letting $I^* = g(T)$ represent this reaction function and recalling that $\hat{T} = 0$, we have private optimization as shown in Figure 5.

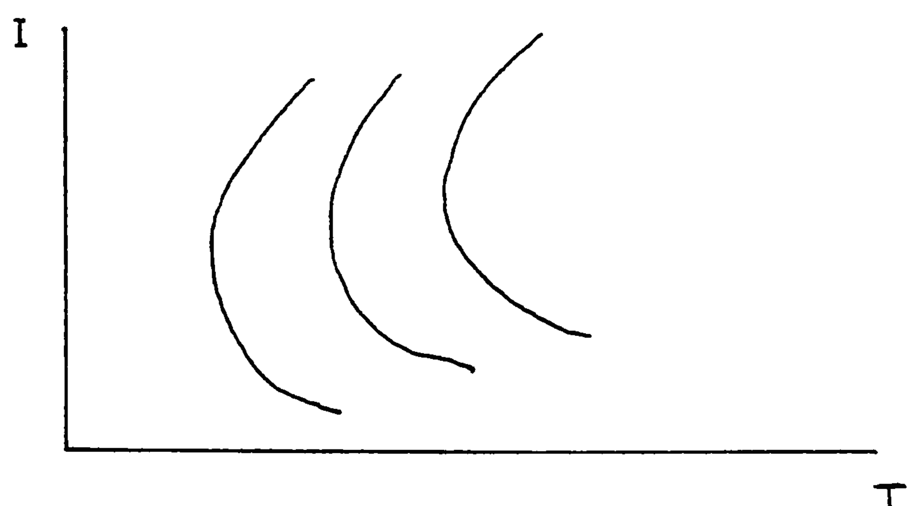


FIG. 4

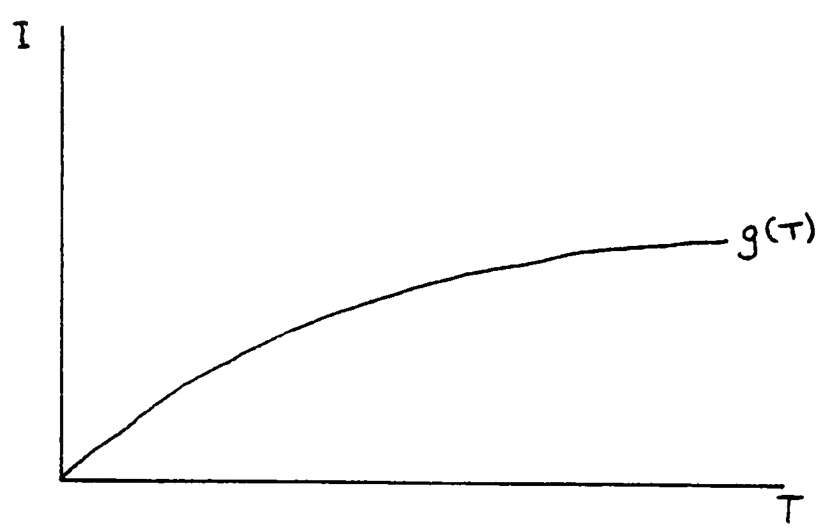


FIG. 5

Government and Social Welfare

The government enters as the second party in the model. The variable under its control is the duration of patent protection. It maximizes a social utility function, U. The government acts as a Stackelberg leader, setting T to maximize U given the knowledge of the private reaction function constraint.

We specify social utility as a function of the net surplus from invention:

$$U = U(S,P)$$

where S is the addition to consumer surplus and P is the addition to producer surplus that results from introduction of the new process. Since both S and P are functions of the level of R&D expenditures and the length of patent protection we can write the indirect utility function

$$U = U[S(I,T),P(I,T)] \tag{7}$$

Working again in I,T space, the social indifference map can be constructed. Following the lead of others, we make the assumption that P and S enter the social utility function linearly. While recognizing the problems inherent in imposition of such a restrictive assumption, we make it merely for ease of exposition. As is pointed out in Appendix A, all results are qualitatively unchanged for any weighting of consumer surplus at least as heavily as producer surplus. If the weight given to P was made stronger, we would eventually reach the point where private and social benefit coincide.

Obviously, as long as patent protection is not infinite in duration, any increase in I will lead to larger consumer surplus effects as we will eventually see a fall in prices due to the improved technology. Additionally, we know that for any level of T there exists some I* which maximizes private discounted profits from invention.

Formally, we have established that, given some T,

$$S_I > 0 \qquad \forall I$$

$$P_I \gtreqless 0 \quad \text{as} \quad I \lesseqgtr I^*$$

$$S(0,T) = P(0,T) = 0.$$

The final line is obvious as S and P are defined as incremental surpluses caused by R&D expenditures, I. Note that this expression implies no indifference curve meets the horizontal (T) axis.

Holding T constant, as we move vertically through I,T space we move through larger iso-utility curves up to the point where $S_I + P_I = 0$; further increases in I result in smaller valued iso-utility curves. To see this, differentiate (7);

$$du = S_I dI + P_I dI + S_T dT + P_T dT$$

Now, if we are moving vertically through (I,T) space, dT = 0. That is,

$$\frac{du}{dI} = (S_I + P_I)$$

which implies increasing utility up till the point where P_I has not only become negative, but exactly offsets the S_I; define this as $I^{**}(T_i)$. On the other hand, given some level of I, we know that social utility is

maximized by a patent period of T = 0. Recall that consumer surplus is producer surplus plus the additional triangle realized when price drops at the expiration of patent protection. This implies that a shortening of T increases consumer surplus by the size of the decrease in producer surplus plus the additional triangle gained. That is,

$$S_T < 0, \quad P_T > 0 \qquad \forall T$$

$$(S_T + P_T) < 0$$

These results combine to give us a social indifference map such as the one pictured in Figure 6. Indifference curves closer to the vertical axis, (i.e., the innermost curves) correspond to higher levels of utility. I_0^{**} is a bliss point for this industry, since even with T = 0, we have an $I < \infty$ which maximizes U(S,P) given the cost of I.

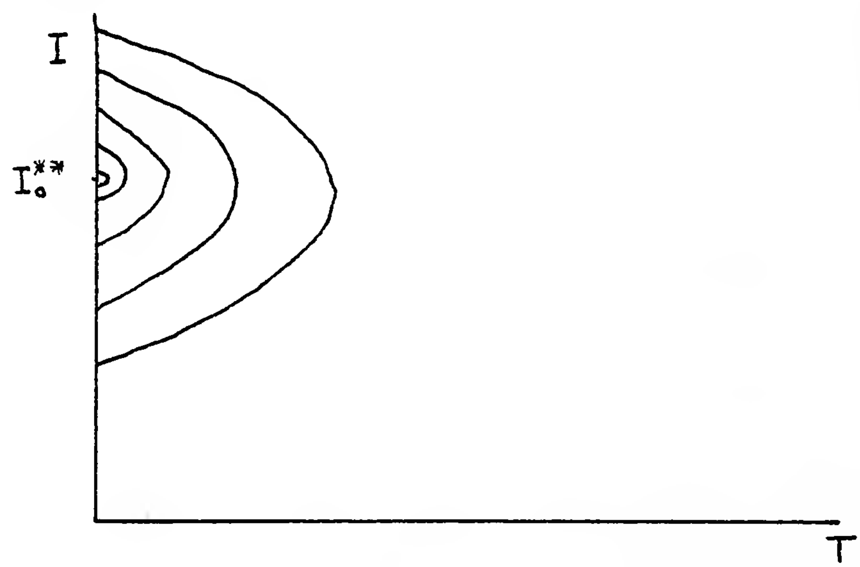


FIG. 6

Mathematically, we can totally differentiate (7) and set dU = 0 to find the slope of the indifference curve as

(8) $$\frac{dI}{dT} = -\frac{U_S S_T + U_P P_T}{U_S S_I + U_P P_I}$$

Again, assuming U_S and U_P are of equal size, the sign of (8) is such that

$$\frac{dI}{dT} > 0 \text{ for } I < I^{**}(T_i)$$

$$\frac{dI}{dT} < 0 \text{ for } I > I^{**}(T_i)$$

where $I^{**}(T_i)$ is defined as the socially optimal level of I for the given T_i

(9) $$S_I(I^{**}(T_i)) + P_I(I^{**}(T_i)) = 0$$

Combining the social indifference map with the private reaction function in I,T space gives the configuration as shown in Figure 7. As is proven in Appendix A, the socially optimal ridge line has negative slope. Additionally, since

$$\begin{array}{ll} P_I > 0 & I < I^*(T_i) \\ P_I = 0 & I = I^*(T_i) \\ P_I > 0 & I < I^*(T_i) \\ S_I > 0 & \forall I \end{array}$$

we know by (9) that $I^{**}(T_i)$ is always greater than $I^*(T_i)$. As T approaches infinity S approaches zero implying that the I** locus and the I* locus asymptotically approach each other.

Acting as the Stackelberg leader, the government can vary patent life policy to maximize the constrained social welfare. The optimal patent life for the case depicted in Figure 7 is T*. However, this

assumed that $\hat{T} = 0$; the only protection available was government patent protection. Now let us relax this assumption.

If $\hat{T}_0 > 0$ then the reaction function g(T) is truncated at the low side by $\hat{T}_0$; intuitively, a patent life less than $\hat{T}_0$ would be reflective as firms in this industry could rely upon natural protection from imitation for a longer period of monopoly returns. The reaction function becomes, in effect, vertical at $\hat{T}_0$ as is shown in Figure 8.

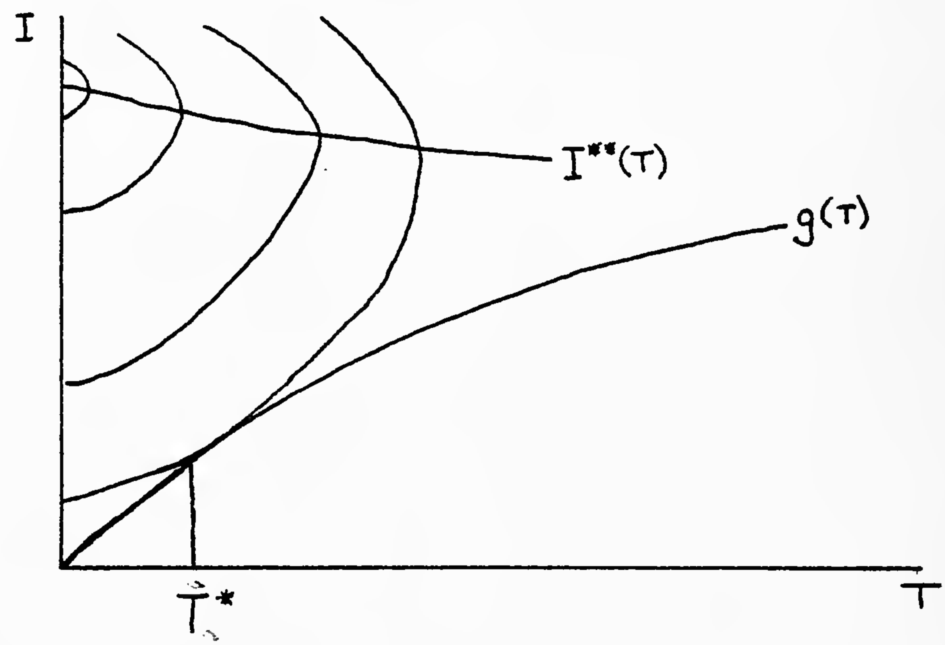


FIG. 7

However, the indifference map was also sensitive to industry conditions and as such will also be altered if $\hat{T} > 0$. Imagine the industry depicted in Figure 7, where $\hat{T} = 0$. Figure 9 shows a representative indifference curve; a locus of (I,T) combinations that gives utility U_0. Now, for any reason, let $\hat{T}_0 > 0$ be the natural protection period. By the same logic as above, this implies that if T is lowered below $\hat{T}_0$, we stay at utility level U_0 only by holding I at I. That is, all indifference curves become horizontal lines between T = 0 and $\hat{T}_0$. It

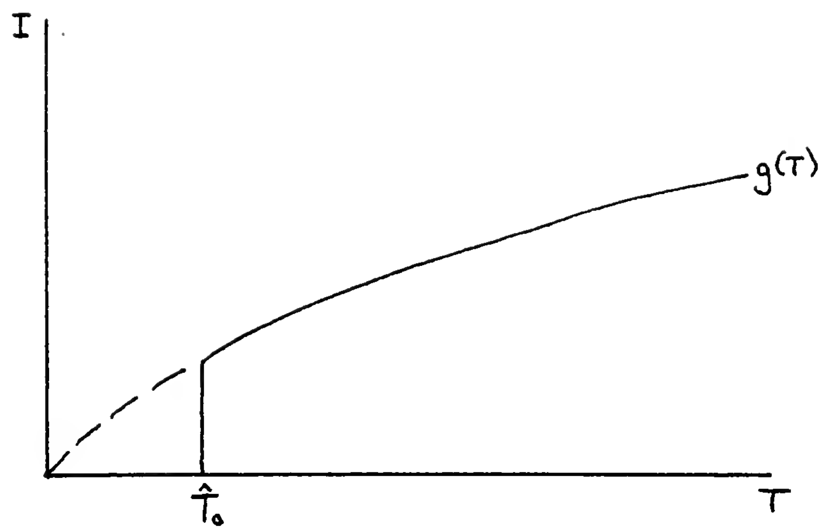
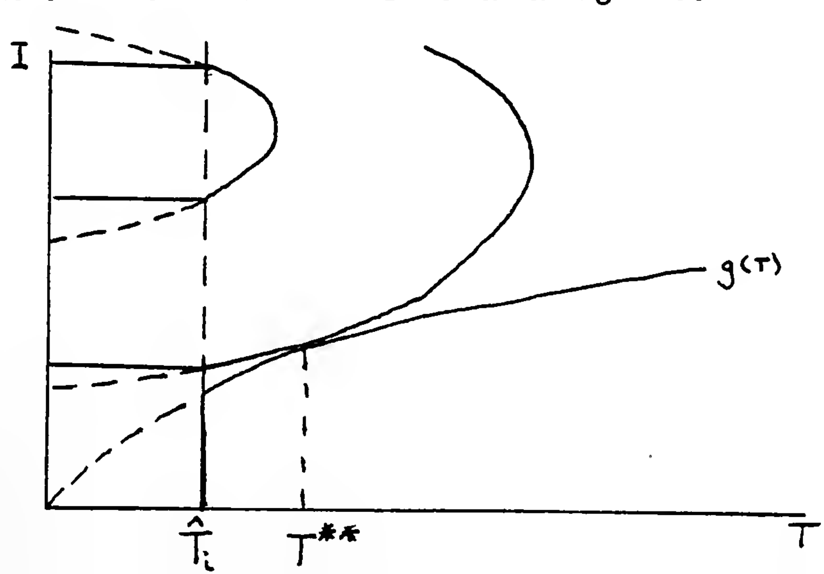


FIG. 8

should be recalled that we are assuming away any external effects that may arise from the fact that patenting increases society's knowledge immediately and this may lead to further opportunities to invent elsewhere.

Finally, some general observations about the effect of $\hat{T} > 0$ on the socially optimum patent life. If the socially optimal T*, determined where the reaction function g(T) is tangent to the innermost indifference curve, is larger than $\hat{T}_i$ (the natural protection period for industry i) nothing is altered. Such a situation is shown in Figure 9.



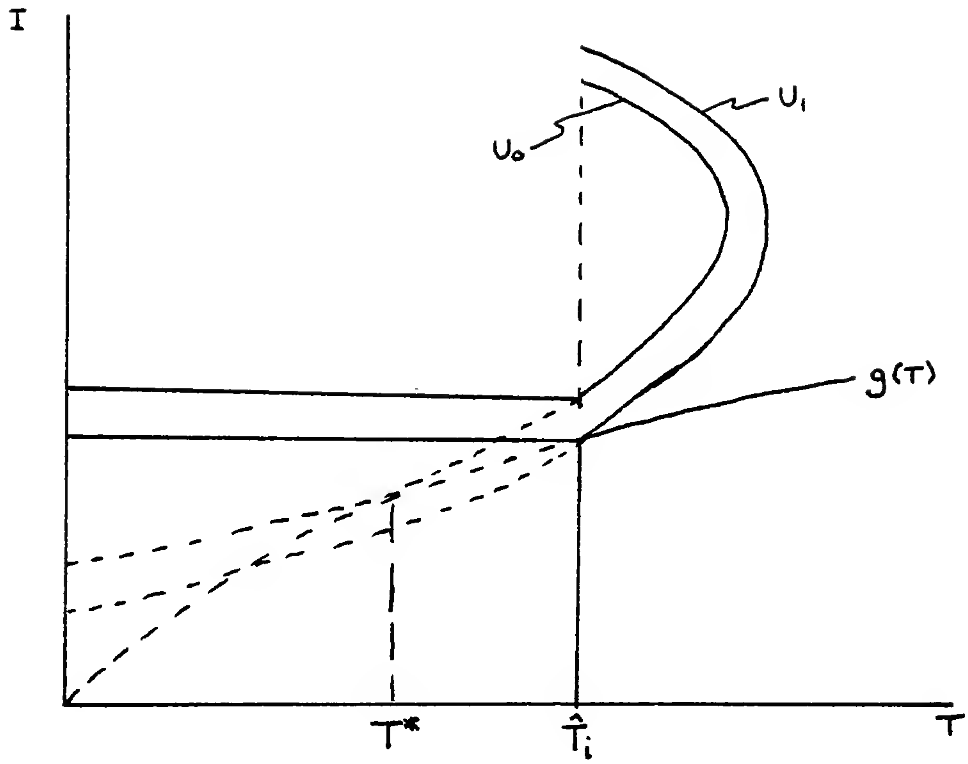


FIG. 10

The dotted lines depict how the curves would appear if $\hat{T}_i = 0$.

On the other hand, some industries have conditions which permit innovators to achieve rents sufficient to induce significant levels of R&D with no patent protection. These industries may have a $\hat{T}_i > T^*$. In this case, a corner solution arises, as is depicted in Figure 10.

The optimal protection period, T^*, is less than what the firm is allowed by natural imitation lags, etc., in its industry. Society made worse off ($U_1 < U_0$) as the price decrease effect is farther away than optimal. At this corner solution, patent life is, effectively, zero for this industry as $\hat{T}_i$ maximizes the constrained welfare.

Finally, you will note that all private reaction functions have become positive in I at T values of $\varepsilon > 0$. In fact, we may have industries where the costs of R&D are large and/or imitation is very very easy. Such a situation is depicted by a very low, perhaps zero, value of $\hat{T}_i$ and a reaction function g(T) that has a horizontal intercept at $\overline{T} > 0$. That is, $g(T) = 0$ for $0 < T \leq \overline{T}$ and $g'(\overline{T}) > 0$. As shown in Figure 11, this alters the indifference map in a way similar to the situation where $\hat{T}_i > T_i^*$. For any level of $T < \overline{T}$, there is no private R&D expenditure forthcoming. Thus, as we vary T_i where $0 < T_i < \overline{T}$ we must hold I constant

if we are to stay at any given level of utility. Thus, the indifference curves, once again, become horizontal. This is due to the specification of utility in the indirect form: a function of consumer and producer surplus. Regardless, the industry in Figure 11 will have an interior solution at a positive patent life.

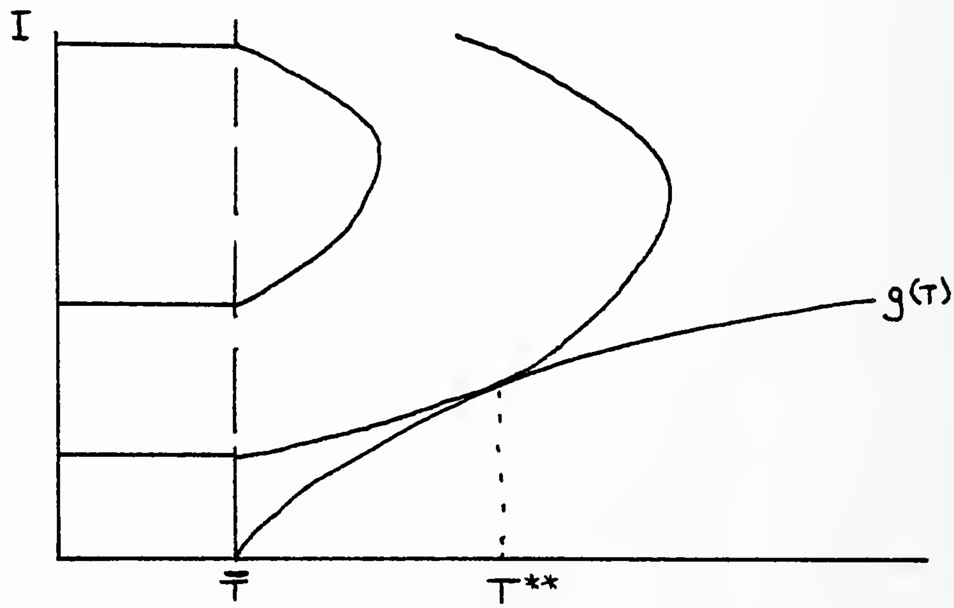


FIG 11

Thus, we see that not all industries need patent protection to bring forth the (second best) optimal R&D projects; this is as has been argued by some previous author's empirical work [see, for example, Scherer (1969)]. Industries such as those depicted in Figure 11 require certain restrictions to allow the invention form to recoup its costs and reward its initiative. On the other hand, Figure 10 showed the case where optimal patent life was zero. Intuitively we can imagine an industry where patent protection is not important; natural protection is longer than the socially optimal patent protection absent any $\hat{T} > 0$.

II. RIVALRY IN THE INVENTION MARKET

In the preceeding analysis, we operated under the assumption of a monopoly inventor. In this section we add more realism by dropping this rather restrictive assumption. In most if not all cases of invention, the recipient of the patent is not the only party to have worked toward gaining patent rights on the specific invention. The winner of the patent will gain monopoly rents which more than offset its expenditure on winning the patent. The losers will get nothing out of their efforts except the loss of expenditures invested in the losing cause. However, the expenditures represent a cost to society which must be considered in our social welfare maximization process.

Kamien and Schwartz (1974) examined the optimal patent life under condition of rivalry. As in this model, the potential inventors recognize the existence of private benefit from a patented invention. Kamien and Schwartz (K-S) specify a concave benefit function as in Nordhaus. Rivalry removes any certainty that the potential inventor will realize the monopoly returns; the expected return to any individual rival will be smaller than if it were a monopoly inventor.

At this point, two specifications in the K-S model become very important. First, in their model the winner of the patent is not allowed to continue development of the invention. Given the lowered expected return brought about by rivalry, the R&D goal of the firm is lower than that of a monopoly developer. By designing the development process such that it halts at the level of R&D which achieves the patent, K-S have made the process putty-clay in nature. It seems clear that once the winning firm has been awarded its patent, it would like to behave like the monopoly

inventor and develop the invention to the actual profit maximizing level. Indeed, Scherer (1969) reports that the majority of expenditures on R&D occur at the development, or innovation, level. Apparently in K-S, the firm ignores the fact that at its winning level of R&D the marginal revenue from an extra dollar of investment is greater than the marginal cost; rivalry diminishes the size of the R&D project and hence the patented final process improvement.

Secondly, the K-S model specifies the degree of rivalry as an exogeneous parameter. The simultaneity problem is obvious; in a model designed to find the optimal duration of patent life under rivalry, one can hardly ignore the effect of patent length on the degree of rivalry itself.

The present model specifies a different inventor-innovation endeavor: a two stage process. The first stage describes the competition to win the patent while the second stage allows the winner to develop the patent to the privately optimal I* just as the monopolist did earlier.

A. The Process of Rivalry

The existence of rivalry will alter the social optimization problem. Rivalry produces duplicative costs in terms of the expenditures made by losers. This rent erosion will cause net producer's surplus to diminish. The exact degree of this rent erosion will be established in what follows. The presence of entry (absent in the K-S formulation) is a key factor in determination of the amount of rent erosion.

Assume that firms attempting to gain a patent expend some amount to enter the patentee selection process.[5] Only one firm can win the patent but many firms can enter the "lottery" by buying a "ticket." For example,

such tickets may be the salary of bright young physicists. The research procedure is assumed subject to some stochastic variation; one firm will win the patent but we do not know a priori which firm. Each firm picks a method toward achieving the level of invention that is sufficient to be awarded a patent. There are many such paths (tickets), each of which are independent and have the same expected value and cost. The cost is given as some constant, γ, which is the sunk cost of proceeding to the goal of a patentable development.[6] We assume a winner emerges early; γ is a small percentage of the final level of R&D chosen by the winner. All losers are denied protection and thus lose the entirety of their investment, γ. After the winner has been chosen, it goes on in the second stage, spending the additional $(I^* - \gamma)$ to arrive at the profit maximizing level of investment in R&D.[7]

The equilibrium condition in the invention innovation market requires that the discounted value of expected profits equals zero; if discounted expected profits are greater than zero, entry into the market for invention will occur. All firms are assumed to know the profit accruing to a winner and the number of firms striving to win. Since we assume that all paths toward winning are of equal expected value, all potential inventors are identical.[8] We can then write the probability that firm i will be the successful inventor as

$$\rho = 1/n$$

where n is the number of firms in the competition for the patent.[9] Each firm realizes that it will expend the additional $(I^* - \gamma)$ if it wins the patent. Discounted expected profits can now be written as

(10) $\pi_e = \rho R - \rho(I^* - \gamma) - \gamma$

where, as before, R is the monopoly inventors discounted rents and the $(I^* - \gamma)$ is the extra I the winning firm will have to make past γ if it wishes to maximize profits.

The equilibrium condition can now be written as (10) = 0; rewriting this gives

(11) $\pi_e = \rho(R - (I^* - \gamma)) - \gamma$.

If π_e is greater than zero, entry into the inventive process will occur, increasing n and lowering ρ. The value of R and I* will be unaffected by the degree of rivalry as they are functions of the extent of patent protection; regardless of the size of n, the winner of the patent will choose the I* and get the same R as in the monopoly model (given the same T). The degree of rivalry is determined by solving (11) for the equilibrium number of firms, n*. Knowing n* it is possible to evaluate the size of duplicative R&D expenditures. That is, the rent erosion due to rivalry is $(n^* - 1)\gamma$, the total value of losing attempts. Recalling that producer surplus P is simply the present discounted value of the monopoly inventor's profits, R - I*, we have from (11)

$$0 = \frac{1}{n^*} (R - I^*) + \frac{1}{n^*} \gamma - \gamma$$

which reduces to

(12) $P = (n^* - 1)\gamma$.

At the zero profit equilibrium, the producer surplus earned by the winner of the patent is entirely eroded by the losers' expenditures in

attempting to gain the patent.[10] The effect of such rivalry will be to lower the social welfare associated with any given (I*,T) combination as net producer surplus is eroded to zero. As in Section I we will first examine a simple model, turning later to a more generalized version. Additionally, we begin the analysis assuming that $\hat{T}_i = 0$.

B. A Simple Model of Rivalry in Invention

We use the same model as described in Part A of Section I, with the notable difference of the addition of the rivalry process described above; consumer and producer surplus are represented as the areas under the same curves as in model I. However, the rent erosion brought on by rivalry will cause net producer surplus to vanish. Given this rent

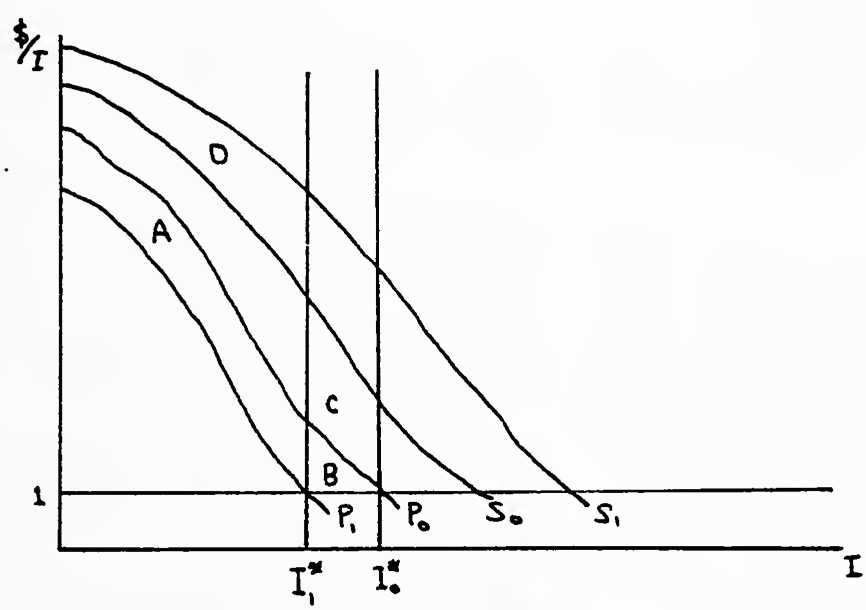


Fig 3; repeated

erosion, we can return to Figure 3 and examine the question of optimal patent life. Suppose the government decreased the optimal patent life. Once again, the P curve shifts down while the S curve shifts out. The positive change in consumer surplus is D + A - C. On the private side,

the winner of the patent will lose the area A + B but society sees a simultaneous drop in resources spent in rent erosion equal to the area A + B. Thus there is a zero net effect on the private side. The change in welfare from a drop in T is now A + D - C. If this is positive, the decrease in T was in the best interest of society.

Comparing this case with the monopoly case, assume we are currently at T_0 in Figure 3 which is optimal given no rivalry. This implies D = C + B. If rivalry were present, a decrease in T_0 would result in a change in welfare equal to A + D - C. However, by the optimality of T_0 in the monopoly case this is equal to

$$(A + C + B - C) = (A + B) > 0.$$

The conclusion is obvious; in a market with rivalry, the optimal patent protection is shorter than when no rivalry exists.

C. A More General Model With Rivalry

Incorporating rivalry into the generalized format presented in Part B, Section I is not difficult with respect to the private sector. Rivalry does not effect the reaction function used earlier; once the patent has been conferred, the winner acts exactly as the monopoly inventor in the previous model. Examination of social utility is somewhat more difficult.

Under conditions of free entry into the inventive process, we saw that rent erosion offsets the patent holder's profit leaving net producer surplus at zero. In equilibrium, we know that social welfare will collapse to the value of consumer surplus. However, this result holds only when I = I*; that is, when we are on the reaction function constraint. In constructing an indifference map, we must examine points other than those

on the reaction function, $I^* = g(T)$. Recalling that all rivals are aware of $g(T)$ and can thus predict the I^* and P associated with winning, the size of duplicative research expenditures are known in advance as $(n^* - 1)\gamma$. Once the government sets any level of T, the rivalry process drives duplicative effort to this level. If the winning firm's eventual R&D is anything other than I^*, its rents will be less than P, implying negative net producer surplus for society.

Once again, consider the situation where social utility is the additive form of net surplus. Let $U = U(S,P)$ denote the social utility function in the monopoly inventor model, while $W = W(S,P,E)$ social utility in the rivalry case where rent erosion (E) is present. W differs from U by the extent of duplicative R&D. Recalling that the loser's expenditures are based entirely upon the optimal $I^* = g(T)$ (more accurately, upon the present discounted value of the profit earned from I^*), we can write the duplicative costs, E, as

$$(13) \qquad E = E(g(T)) \qquad E' > 0$$

Recalling the assumption of additivity, this gives:

$$(14) \qquad W(S,P,E) = U(S,P) - E(g(T))$$

Several possible cases arise, depending upon the size of $\hat{T}_1$, the natural protection lag. All but the last case result in the same qualitative conclusions but each present slightly different indifference maps.

<u>Case 1</u> $\hat{T}_1 = 0$. In this case, there exists no natural imitation protection; the private reaction function is not truncated. Begin by examining the case where the reaction function starts at the origin;

$g(0) = 0$ and $g'(0) > 0$. Given any $T > 0$ we have $I^* > 0$ which means that rivalry costs will push $W(I_i,T_j) < U(I_i,T_j)$. Differentiating (14), setting $dU = 0$, and solving for dI/dT we can examine the slope of the indifference curves under rivalry as

$$(15) \qquad \frac{dI}{dT} = - \frac{S_T + P_T - E'g'}{S_I + P_I}$$

The denominator is identical to that found in the monopoly model for any given I,T combination (see equation 8). Therefore the indifference curves become vertical at the same I^{**} for any given T; the ridge lines for the monopoly and the rivalry case are identical. Since $E'g'$ is a positive term, we know the numerator is still everywhere negative as in the earlier model. Thus the indifference curves in the rivalry framework will change slope along the same I^{**} line as found earlier.

Additionally, since $E'g' > 0$, the numerator of (15) is greater in absolute value than the numerator of (8). This insures that at any (I_i,T_i) combination where $I \neq I^{**}(T_i)$ the slope of the indifference curve associated with W will be steeper than those of U.

Finally, we can picture representative indifference curves by recalling that $g(0) = 0$ and $g'(0) = 0$. This implies that at $T = 0$ there are no duplicative R&D expenditures and $W = U$. In Figure 12, the indifference curve corresponding to W becomes vertical at the same $I^{**}(T_i)$ as its counterpart representing U. However, the indifference curve for the U function is flatter than that for the W function at every other (I,T) combination. Additionally, we know that the $\overline{U}_i$ indifference curve represents higher social welfare than the $\overline{W}_j$ shown in Figure 12. At any T_i chosen by the government, the vertical slope

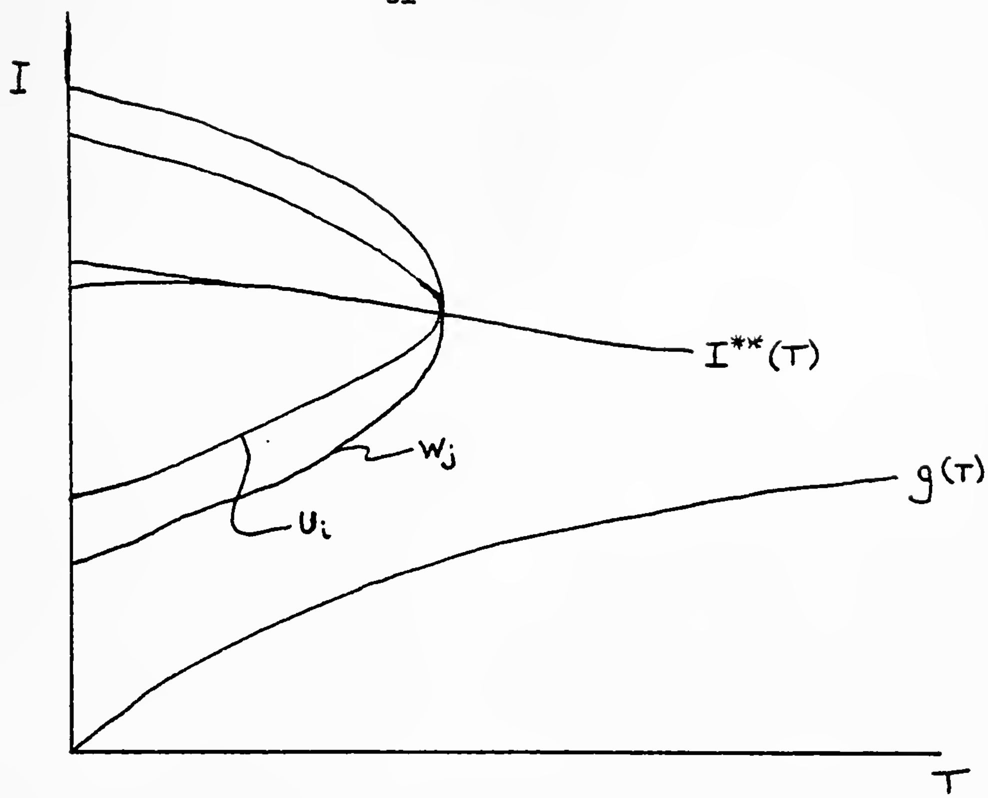


FIG. 12

point, $I^{**}(T_i)$, is the same on both indifference maps. However, $W < U$ at this point due to rent erosion; to keep the same level of social welfare given some I, the government must shorten the period of monopoly protection. Now examine the situation where $\hat{T} = 0$, but $g(T_i) = 0$, $0 < T_i < \overline{T}$. This refers to the case where the reaction function has a horizontal intercept at point $\overline{T}$. The analysis is similar to above except as shown earlier in the monopoly model, the indifference curves are horizontal for any T less than $\overline{T}$. Given some I, varying T between zero and $\overline{T}$ brings forth no extra benefit or cost whatsoever; utility remains constant. Note also that this level of utility is the same for both models, as no rivalry costs have come into play. However, for any $T > \overline{T}$, the reaction function gives $I^* > 0$ which implies rivalry in the inventive process. At such a $T_0 > \overline{T}$ the level of I will need to be greater in the case of rivalry in order to keep the

level of W equal to that of U. The indifference curves for W are shown are always steeper than those for U except at $I^{**}(T_i)$. The indifference map for case 2 is shown in Figure 13.

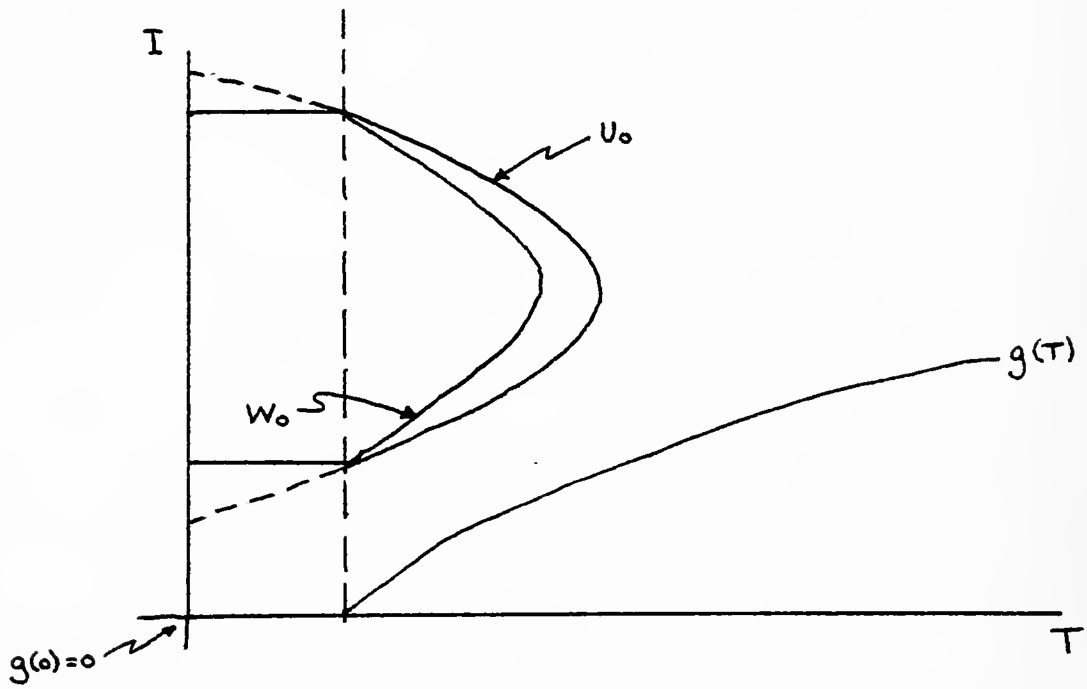


FIG.13

Recognizing that the government should act according to the reaction function constraint, we should always be operating on indifference curves not coincident with those derived in Section I; Figure 12 gives a qualitatively representative indifference map of the $\hat{T}_1 = 0$ case. The government's optimization problem is the same as in the first section. However, it is obvious the optimal T in any given market will be lower in the case of rivalry. Figure 14 shows the reaction function g(T) for some market, a function unaltered by the degree of rivalry. The optimal patent life in the monopoly model is point E where g(T) is tangent to the highest indifference curve corresponding to U(S,P). We have demonstrated above that at this (I,T) combination, the indifference curve corresponding to W is steeper than that derived from U. The tangency of g(T) occurs at

a point B in the rivalry case, necessarily pointing to a lower optimal T for the Stackelberg leader. Whereas earlier models have found optimal patent life to be increasing in the degree of rivalry, we find the opposite.

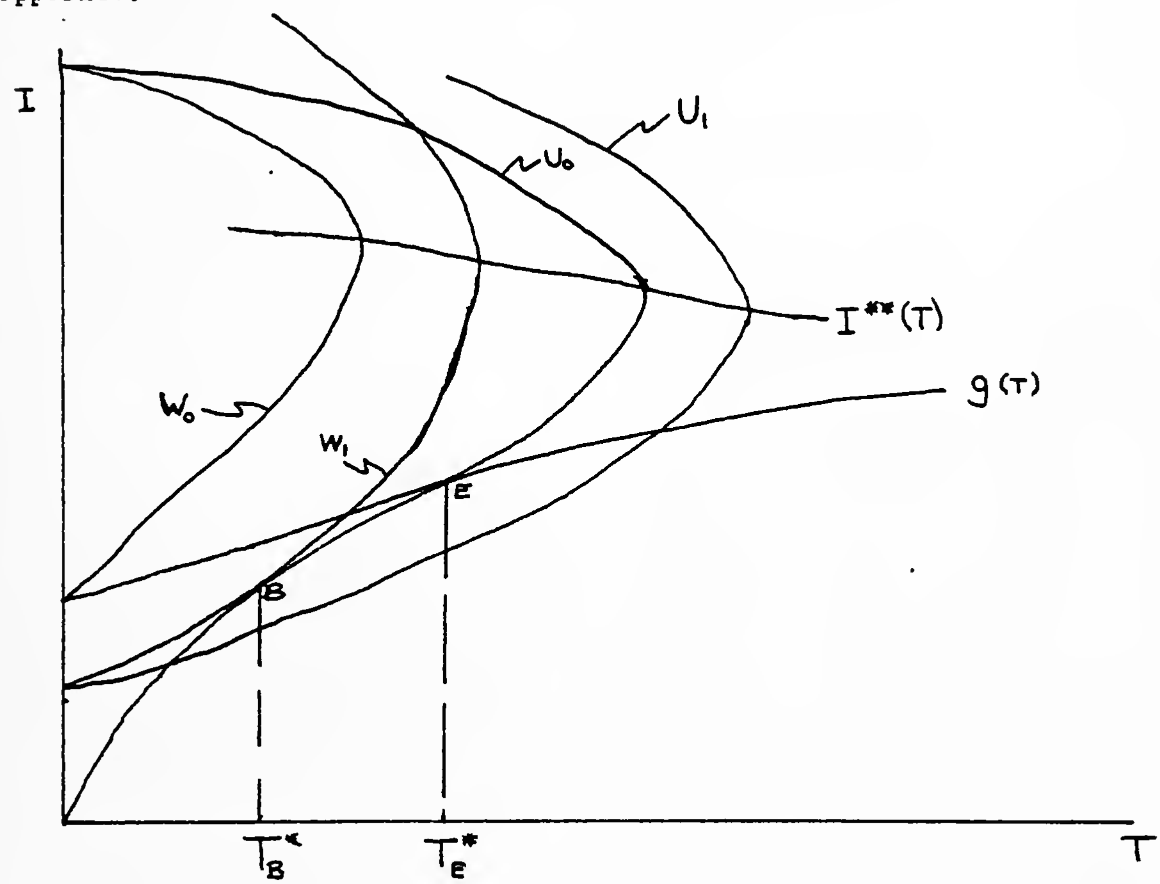


FIG. 14

<u>Case 2</u>: $0 < \hat{T}_i < T^*_m$ where T^*_m is the optimal T under the monopoly inventor assumption. Once again, we will work in the case where the reaction function becomes positive at $T = \varepsilon > 0$. The situation where $g(T) = 0$ for $0 < T < \overline{T}$ follows according to the same logic as above. The existence of $\hat{T}_i > 0$ truncates the g(T) at $\hat{T}_i$. However, since $T^*_m > \hat{T}_i$, we have a positive patent life such as was depicted in Figure 9. The existence of a positive level of natural protection will alter the indifference map as follows. To begin with examine the configuration in

Figure 15, where $\hat{T}_i$ is temporarily held at zero. Now, let $\hat{T}_i$ become the positive value X. This implies that altering patent life between

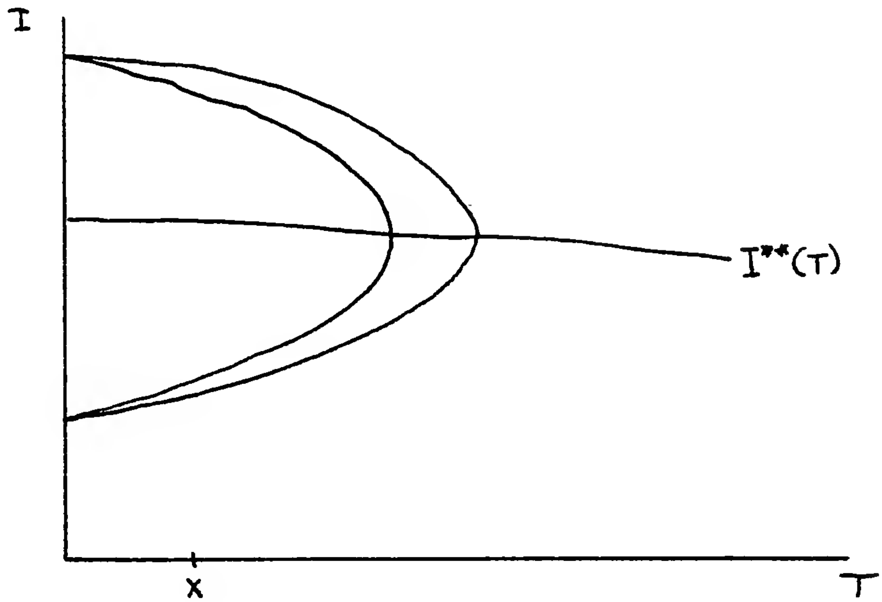


FIG 15

$0 < T < X$ will have not effect on the values of S and P; the indifference curve corresponding to the monopoly situation (i.e., U(S,P)) will become horizontal at the $\hat{T}_i$ constraint. Figure 16 shows the indifference

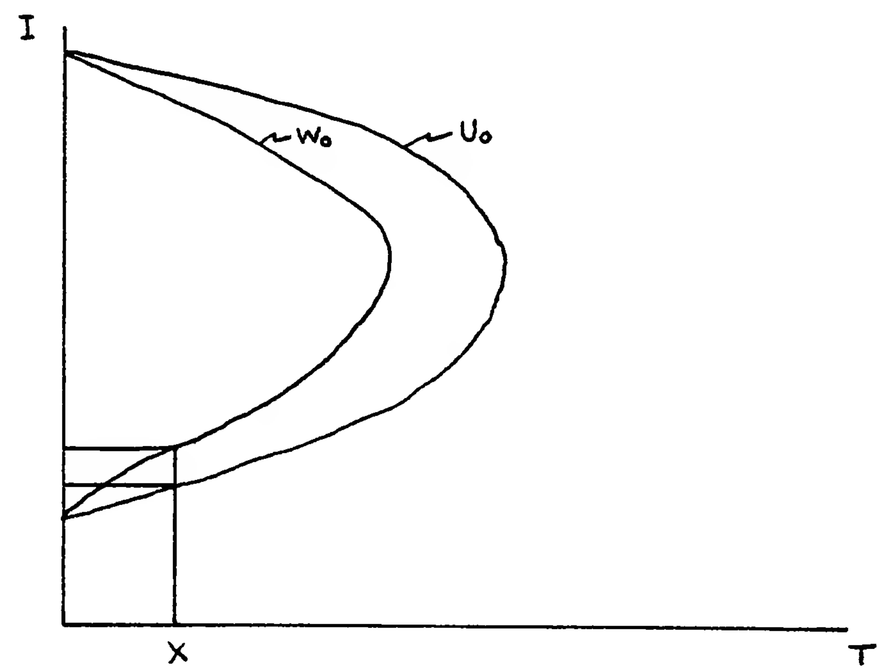


FIG. 16

curve for U_0 now that $\hat{T}_i = X$. Moving to the rivalry case, the indifference curve associated with $W_0(=U_0)$ also becomes horizontal at $\hat{T}_i = X$. At the point (I_0,X), the level of social welfare associated with the (WS,P,E) function is less than in monopoly. Regardless of the source, protection of X exists and rivalry will lead to some losing expenditures. To attain a $W_0 = U_0$ given X is the level of protection, I, must be the level of investment by the winning firm.

Recalling that $0 < \hat{T}_i < T^*_m$ in the present case, we are assumed that T^*_r is strictly less than T^*_m. Figure (17a) shows the case where

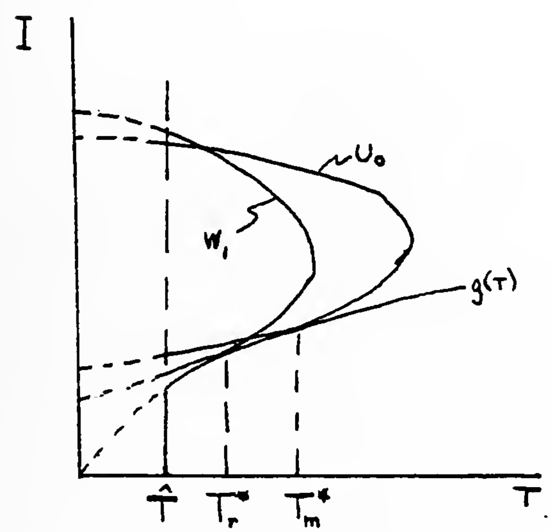


FIG 17a

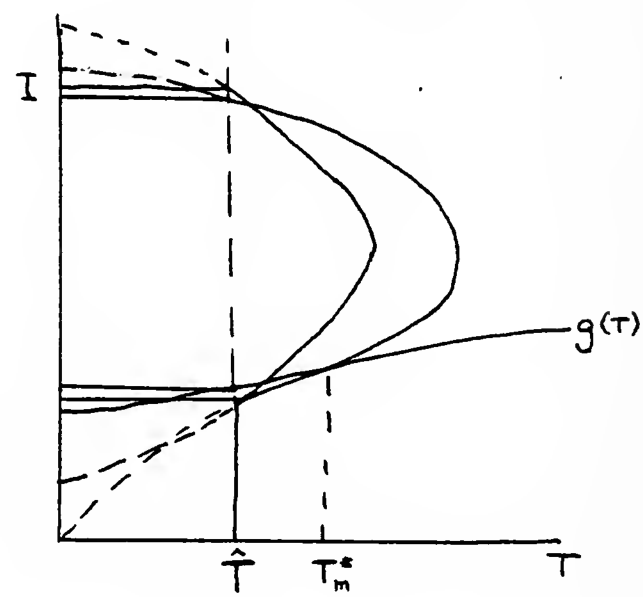


FIG 17b

there is an interior tangency between g(T) and an indifference curves in the W(S,P,E) indifference map. The other possible case is that the optimum is the corner solution where no patent protection is best; social welfare is maximized where the truncated g(T) touches the highest indifference curve but

(16) $g'(T) < MRS_W$

This situation is depicted in Figure (16b). Once again, as in case 1 we can be certain that the optimal patent life in rivalry is at less than in the case of a monopoly inventor.

Case 3: $\hat{T}_i > T_m^*$. This final case is represented in Figure 10. Here, optimal patent life is zero since its unconstrained value of T_m^* is less than the natural protection, rendering it ineffective. We are at a corner solution in both the rivalry and the monopoly case, as shown by Figure 18.

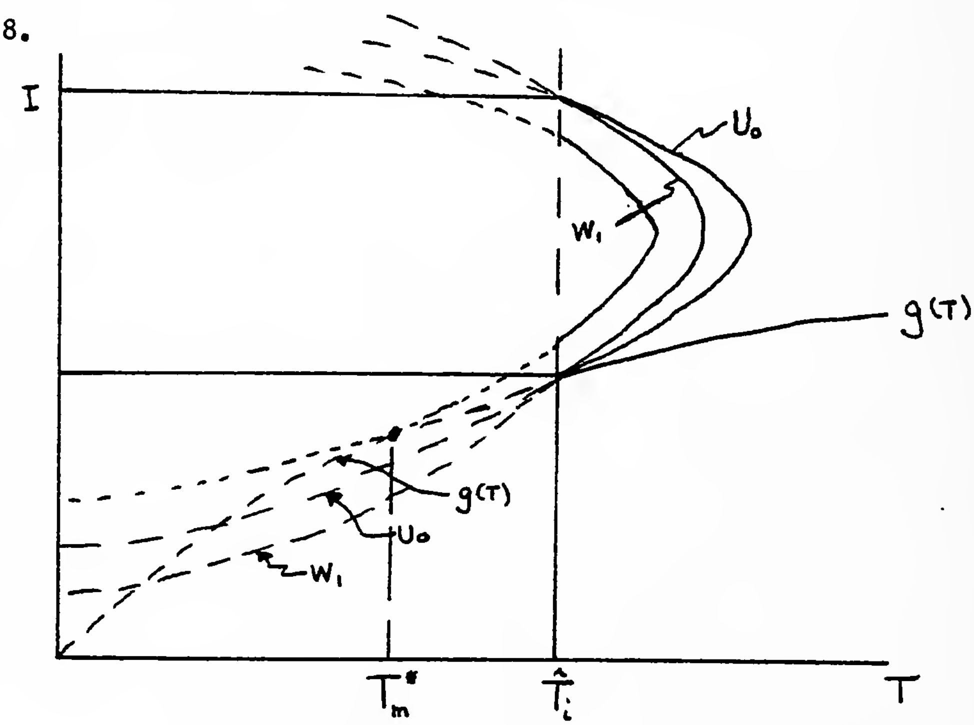


FIG. 18

At this corner solution, $(I_0, \hat{T}_i)$ the optimal patent life under either market structure is zero. Thus, only in this final case is T_r^* as large as T_m^*. But this is only by default as case 3 represents an industry where natural protection is so large that no patent protection is ever warranted.

III. CONCLUSIONS

Using a generalized reaction format, we have shown that rivalry in the invention market leads to duplicative research costs not present in the monopoly inventor model. Rivalry will become more intense, i.e., the number of firms entering the race for the patent will increase, until the present value of expected profit is driven to zero. Under such conditions a different, lower, optimal patent life is arrived at than what resulted from the analysis of the monopoly model. In their study of rivalry's effect on optimal patent life, Kamien and Schwartz found the opposite: rivalry points to a longer optimal patent life. This result follows from their "putty-clay" representation of the inventive process and the exogenous nature of rivalry. Rivalry drives the size of the R&D project down and the winner does not continue development. To counter this, their model calls for a longer patent protection to bring forth larger project size; longer patent protection implies higher profits to the winner but the intensity of rivalry, and therefore the individual firms probability of winning, is unchanged due to its exogenous nature. By specifying rivalry in a more realistic manner, we arrive at a shorter patent life in the rivalry model than in the monopoly model.

Finally a caveat must be noted. In the K-S model, timing of innovation is a factor ambiguously affected by rivalry, whereas our model has not examined this issue. Our results can, at the very least, be considered important if for no other reason than their indication of the sensitivity of optimal patent life to variables that K-S do not examine. Arguments that the present study is biased as a result of

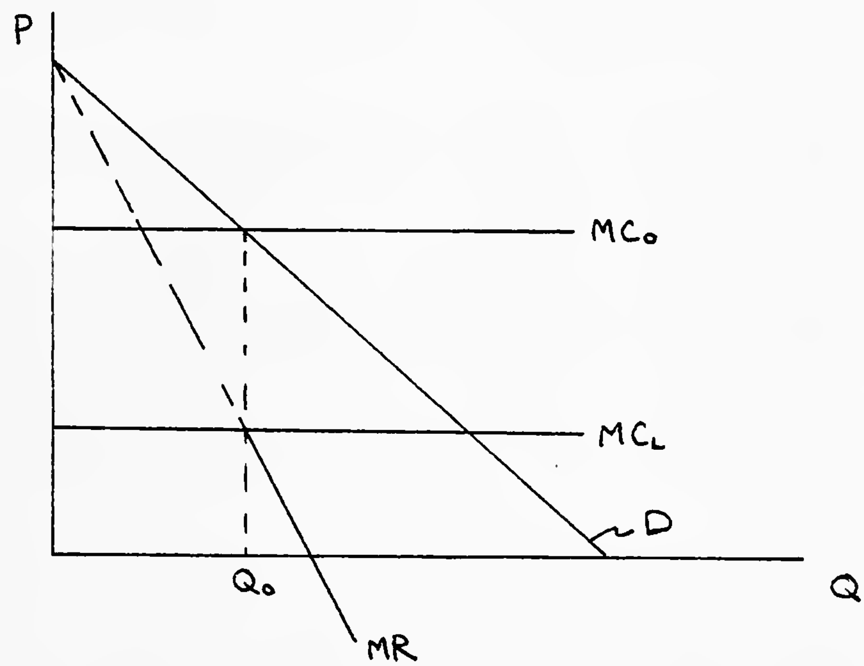
ignorance of timing problems are no more important than arguments that the Kamien and Schwartz model is biased due to the exogenous specification of rivalry.

Additionally some concern over the assumption of certainty of eventual invention may arise. It may indeed be the case that the probability of eventual invention by some firm is an increasing function of the degree of rivalry. However, Ippolito (page 17) has shown that, in general, rivalry implies more social costs through its diminishing effects on project size and its duplicative costs than the possible social benefits it provides through increased probability of invention. This lends little weight to the argument that patent life would be longer in rivalry as it draws forth more rivalry. In a specification similar to the model of this essay, Ippolito specified all the uncertainty prior to the patenting and found that rivalry adds only a marginal degree of certainty to the expected income. With this in mind, it seems any bias introduced from lack of modeling a certainty effect due to rivalry is very small.

FOOTNOTES

1. The qualitative results are not sensitive to this assumption. Monopoly power rather than competition in the given final product market would lower the benefits of any given size process improvement. However, the benefits are still there and hold regardless of the market structure in the invention market.

2. As mentioned above, we assume that MC_1 is not so radically lower than MC_0 as to cause the patent monopolist to desire some $Q > Q_0$. This means that MC_1 must fall between MC_0 and MC_L in the picture below since residual MR faced by the patent monopolist is the line ABCE. Again, violation of this assumption would allow increased benefits; the model would be the same except for a more complicated formulation.



3. Note that the exogenous development lag has been suppressed as it offers no qualitative differences. We may imagine that development compression is very inexpensive so development is reduced to $\varepsilon \to 0$.

4. It is not difficult to show that licensing is almost always the most profitable strategy for a holder of a process improvement patent in a competitive product market.

5. In modeling the process of rivalry, we draw heavily upon the ideas presented in Richard Ippolito's working paper.

6. γ need not be a constant; the cost of entering the "race" could be modeled as θI^*, a cost proportional to the eventual size of

the R&D project. Such a change does not alter the qualitative conclusions, only the effects of changes in expected profit on the equilibrium number of firms.

7. It may be that γ is not the cost minimizing expenditure to reach the patentable project size due to increasing compression costs, etc. However, it can still be the same for all firms, such as is the case in Loury (1975). If γ is not cost minimizing, then the additional expenditure may be $(I^* - \gamma^*)$ where $0 < \gamma^* < \gamma$. As will become obvious, this has no qualitative effects on the results as long as the γ, γ^* difference is known and common.

8. If a firm undertakes more than one path toward inventing the goal process, it will increase its probability of being the winner. However, it will also increase its costs in a manner which makes one firm undertaking j paths qualitatively identical to j firms undertaking one path; there are no economies of scale nor externalities in seeking the patent.

9. As stated above, we assume the probability of someone winning the patent is one. If, however, the probability is something less than one, say ϕ, then $\rho = \phi/n$. This does not alter the qualitative impact of the model.

10. If, as in footnote 7, we have $0 < \gamma^* < \gamma$, the rent erosion is still complete. To see this, suppose that $\gamma = \gamma^* + b$. Expected profit will then be

$$\pi_e = \rho[R - (I^* - \gamma^*) - \gamma] - (1 - \rho)\gamma$$

or

$$\pi_e = \rho[R - I^* - b] - (1 - \rho)\gamma$$

Finally, recalling that $\rho = \frac{1}{n^*}$ and that $R - I^* - b$ is now the winning firms profits, where b is the compression cost payment, we have the $\pi_e = 0$ equilibrium condition

$$R - I^* - b = (n^* - 1)\gamma$$

or

$$P = (n^* - 1)\gamma$$

APPENDIX A

1. Private Reaction Function.

From the profit equation, we know that MR = MC at maximum profit. That is,

$$R_I - C_I = 0. \tag{A.1}$$

Differentiating this with respect to T we will be able to determine the slope of the reaction function: the way privately optimal I changes when T changes.

$$\frac{d[R_I - C_I]}{dT} = R_{II}\frac{dI}{dT} - C_{II}\frac{dI}{dT} + R_{IT} - C_{IT} = 0$$

or

$$\frac{dI^*}{dT} = \frac{C_{IT} - R_{IT}}{R_{II} - C_{II}} \tag{A.2}$$

From the second order conditions for profit maximization, the denominator is negative. Since the level of patent protection does not affect marginal R&D cost, $C_{IT} = 0$. In addition, we know that as T increases, R_I must increase as for any I the period of monopoly protection is longer: $R_{IT} > 0$. An increase in T will cause the marginal revenue function to increase for any level of I, as the effects of a change in I are now felt over a longer period of time. This gives the sign of (A.2) as

$$\frac{dI^*}{dT} = \frac{-(+)}{-} > 0.$$

The firm's reaction function is positive in slope for all $0 < T < \infty$.

REFERENCES

Arrow, Kenneth, "Economic Welfare and the Allocation of Resources for Invention," in Nelson, The Rate and Direction of Inventive Activity, Princeton: Princeton University Press, 1962, pp. 609-625.

Ippolito, Richard A., "A Theoretical Evaluation of Alleged Effects of Patent Rivalry," unpublished paper, April 1976.

Kamien, Morton I. and Nancy L. Schwartz, "Market Structure and Innovation: A Survey," Journal of Economic Literature, March 1975, 13(1), pp. 1-37.

__________, "Patent Life and R&D Rivalry," Northwestern University Working Paper No. 121-72; 1972.

__________, "Patent Life and R&D Rivalry," American Economic Review, March 1974, Vol. 64(1), pp. 183-87.

Loury, G., "Market Structure and Innovation," Northwestern University Working Paper, 1975.

Machlup, Fritz, The Production and Distribution of Knowledge in the United States, Princeton: Princeton University Press, 1962.

Nordhaus, William D., Invention, Growth and Welfare, Cambridge: MIT Press, 1969.

Scherer, F. M., Industrial Market Structure and Economic Performance, Chicago: Rand McNally, 1969.

__________, "Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation," American Economic Review, June 1972, 62, pp. 422-27.

Schmookler, Jacob, Invention and Economic Growth, Cambridge, Mass.: Harvard University Press, 1966.

2. Social Indifference Curves.

From (8) we have the slope of the indifference curve as

$$\frac{dI}{dT} = - \frac{U_S S_T + U_P P_T}{U_S S_I + U_P P_I} \tag{A.3}$$

Initially, we assume that $U_S = U_P$ for ease of exposition; as will be pointed out later, this is not necessary for our results.

Given $U_S = U_P$ we can evaluate the sign of (A.3) as

$$- \frac{(S_T + P_T)}{S_I + P_I} .$$

From the text, we know that $(S_T + P_T)$ is always less than zero. Turning to the denominator, we know that S_I is always positive for $T < \infty$ and $P_I \gtrless$ as $I \lessgtr I^*$. However, as I is increased past I*, the negative P_I eventually offsets the positive S_I. We know this occurs at I** where I** is defined as

$$S_I(I^{**}) + P_I(I^{**}) = 0.$$

Thus, the sign of A.3 can be written as

$$\text{sign}\ \frac{dI}{dT} = \begin{cases} -\left(\frac{-}{+}\right) > 0 & \text{for} \quad I < I^{**} \\ -\left(\frac{-}{-}\right) < 0 & \text{for} \quad I > I^{**} \end{cases} .$$

It should be noted that as long as $U_S \geq U_P$, this result qualitatively is unchanged. If, for example, U_S carried a higher weight, then society would prefer to see consumers receive more of the benefit of invention; that is,

a lower T should be predicted. In fact, $U_S > U_P$ leads to steeper sloped indifference curves as seen in (A.3) which will give just such a result.

3. <u>The Slope of the Ridge Line</u>: I** locus.

We know from (A.3) that the locus of points that make up the $I^{**}(T_i)$ line is the locus of (I,T) combinations where $S_I + P_I = 0$. Totally differentiating this implicit function and solving for dI/dT we will find the slope of the ridge line:

$$\frac{d(S_I + P_I = 0)}{dT} = (S_{II} + P_{II}) \frac{dI}{dT} + S_{IT} + P_{IT} = 0 \qquad \text{(A.4)}$$

Rearranging (A.4) gives

$$\frac{dI^{**}}{dT} = - \frac{(S_{IT} + P_{IT})}{(S_{II} + P_{II})} \,. \qquad \text{(A.5)}$$

Looking first at the denominator, we know that both S_{II} and P_{II} are negative due to the concavity of the benefit function. In the numerator, we know that $P_{IT} > 0$ and $S_{IT} < 0$. At the margin, a decrease in T gives an increase in S equal to the loss in P plus the deadweight loss now retrievable. Thus, the absolute value of the P_{IT} will be less than the absolute value of S_{IT}. That is,

$$(S_{IT} + P_{IT}) < 0.$$

Substituting this into (A.5) we get the sign of the slope of the I** locus as

$$\frac{dI^{**}(T)}{dT} = - \left(\frac{-}{-}\right) < 0.$$